

First Results of ν_e Appearance Analysis and Electron
Neutrino Identification at NOvAJIANMING BIAN
(FOR THE NOvA COLLABORATION)*School of Physics and Astronomy
University of Minnesota, MN 55455, USA*

NOvA is a long-baseline accelerator-based neutrino oscillation experiment that is optimized for $\nu_\mu \rightarrow \nu_e$ measurements. It uses the upgraded NuMI beam from Fermilab and measures electron-neutrino appearance and muon-neutrino disappearance at its Far Detector in Ash River, Minnesota. The ν_e appearance analysis at NOvA aims to resolve the neutrino mass hierarchy problem and to constrain the CP-violating phase. The first data set of 2.74×10^{20} protons on target (POT) equivalent exposure taken by NOvA has been analyzed. The first measurement of electron-neutrino appearance in NOvA provides solid evidence of $\nu_\mu \rightarrow \nu_e$ oscillation with the NuMI beam line. Electron-neutrino identification is the key ingredient for the ν_e appearance analysis. The electron-identification algorithm used to produce the primary results presented here compares 3-D shower-energy profiles with Monte Carlo prototypes to construct likelihoods for each particle hypothesis. Particle likelihoods, among other event-topology variables, are used as inputs to an Artificial Neural Network for the final electron-neutrino identification. The design and implementation of this algorithm is also presented.

PRESENTED AT

DPF 2015

The Meeting of the American Physical Society
Division of Particles and Fields
Ann Arbor, Michigan, August 4–8, 2015

1 Introduction

NOvA (NuMI Off-Axis ν_e Appearance Experiment) is a neutrino experiment optimized to observe the oscillation of muon neutrinos to electron-neutrinos [1]. NOvA uses a 14-kt liquid scintillator Far Detector (FD) in Ash river, Minnesota to detect the oscillated NuMI (Neutrinos at the Main Injector) muon neutrino beam produced 810 km away at Fermilab [2]. The NOvA baseline is the longest in operation, which maximizes the matter effect and allows a measurement of the neutrino mass ordering. NOvA is equipped with a 0.3-kt functionally identical Near Detector (ND) located at Fermilab to measure unoscillated beam neutrinos and estimate backgrounds at the FD. Both detectors are located 14 mrad off-axis to receive a narrow-band neutrino energy spectrum near the energy of the $\nu_\mu \rightarrow \nu_e$ oscillation maximum range (~ 2 GeV), enhancing the $\nu_\mu \rightarrow \nu_e$ oscillation signal in the FD while reducing neutral current and beam ν_e backgrounds from high-energy, unoscillated beam neutrinos.

NOvA's detectors consist of plastic (PVC) extrusions filled with liquid scintillator, with wavelength shifting fibers (WLS) connected to avalanche photodiodes (APDs). The dimensions of the detector cells are 6 cm \times 4 cm, with each cell extending the full width or height of the detector, 15.6 m in the FD and 4.1 m in the ND. Extrusions are assembled in alternating layers of vertical and horizontal extrusions plane, so 3-D hit information is available for clustering and particle identification. Each plane (cell width) of the detectors is just 0.15 radiation lengths (X_0). This level of granularity helps greatly to separate electrons from π^0 backgrounds [3, 4, 5].

NOvA measures ν_e ($\bar{\nu}_e$) appearance probability and ν_μ ($\bar{\nu}_\mu$) disappearance probability with neutrino and anti-neutrino beams. The ν_e ($\bar{\nu}_e$) appearance experiment investigates (1) the neutrino mass hierarchy, (2) the CP-violation phase in the neutrino sector and (3) the θ_{23} octant (whether $\theta_{23} >$ or $< 45^\circ$). For the first ν_e analysis, we use 2.74×10^{20} POT equivalent exposure (1/13th of the overall planned exposure) of neutrino running to measure the probability of $\nu_\mu \rightarrow \nu_e$. Although there is no information from $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ in the first analysis, in a favorable δ_{CP} region around $3\pi/2$, there is some sensitivity to discriminate the mass hierarchy.

The first ν_e appearance measurement is a cut-and-count analysis [6]. A ν_e event is identified in charged current (CC) interactions where the electron-neutrino converts into an electron. The ν_e analysis at NOvA makes use of two electron event identification algorithms (EID). The primary EID algorithm is LID, which is an artificial neural network using 3-D shower-energy profile based likelihoods for particle hypotheses. The second EID is LEM, which matches trial events to an enormous Monte Carlo library. Implementation of the LID algorithm is described in Section 2; detail about the LEM algorithm can be found in Ref. [7].

2 Electron Neutrino Identification (LID)

The basic idea of LID is to use the shower-energy profile to separate electrons from muons, π^0 s, and other hadrons. Different particles have very different energy-depositions in the detector. For example, the electron deposits energy through ionization in the first few planes then starts a shower; the photon is a shower that follows a gap in the first few planes; and the muon registers as a long minimum ionizing particle (MIP) track. This makes it possible to identify particles by comparison of shower/track shapes with different particle hypotheses.

To precisely identify particles, we cluster showers from reconstructed collections of cell hits with a start point and direction (prong), and then compare the measured longitudinal and transverse dE/dx of candidate showers with the expected distributions found in MC samples for each particle hypothesis to obtain likelihoods. We perform this comparison plane-by-plane in the longitudinal direction and cell-by-cell in the transverse direction. In this way we can make use of all energy-profile information in a shower. Using these likelihoods, a neural network has been trained and applied to the identification of ν_e CC events.

2.1 ν_e event reconstruction

The ν_e event reconstruction begins with clustering hits by space-time coincidence to separate beam events from cosmic rays in a trigger window. This procedure can collect together hits from a single neutrino interaction (slice). The slices then serve as the foundation for all later reconstruction stages [8]. Next, a modified Hough transform is applied to identify prominent straight-line features in a slice. Then the lines are tuned in an iterative procedure until they converge to the interaction vertex of that slice. Prongs are then reconstructed based on distances from hits to the lines associated with each of the particle paths emanating from the reconstructed vertex [9]-[13].

2.2 Shower clustering

We define the shower core based on the prong direction provided by the prong cluster, then collect signal hits in a column around this core. To reduce the contamination from the hadronic interaction around the vertex, we require the radius to be twice the cell width for the first 8 planes from the start point of the shower and 20 times the cell width for other planes. In this way, there is good efficiency for including all hits caused by an electron in the shower.

2.3 Cell energy deconvolution

Daughter photons from a high-momentum background π^0 and a shower in a neutrino event could partially overlap. To handle this, we perform a de-convolution when we detect overlapping showers based on total energy or distance to the cores. For a cell associated with more than one shower, we determine the energy associated with the i -th shower as follows:

$$E_i^{cell} = \frac{PE^{cell}}{a_i} \cdot \frac{(PE_i^{shower} P_i/a_i)}{\sum_j (PE_j^{shower} P_j/a_j)},$$

$$P_j = \exp(-D_j/\lambda),$$

where E_i^{cell} is energy in a cell belonging to the shower, PE_i^{shower} is the total number of photoelectrons (PE) in the shower, PE^{cell} is the total PE in a cell, a_i is a factor that scales from PE to GeV and corrects for attenuation based on distance to readout, D_i is the distance from the cell to the core of the shower, and λ is a constant for the shower lateral profile. Here we assume the transverse dE/dx as a function of distance to the shower core behaves as follows:

$$dE/dx(x) = AE \exp(-x/\lambda),$$

where A is the normalization constant, E is the shower energy and x is the distance to the shower core. By fitting the transverse shower energy profile of simulated electrons from ν_e we determine λ to be 3.05 cm.

2.4 dE/dx and log likelihood

For an unidentified particle, we compare the measured dE/dx with the expected dE/dx under each particle hypothesis in each plane and transverse cell to construct the probability and likelihood for particle identification. In this way we can describe the 3-D development of a shower in detail.

Using the deconvoluted cell energy, we calculate dE/dx in the longitudinal and transverse directions. The longitudinal dE/dx is calculated plane-by-plane. It is defined as the total shower energy in a plane divided by the path length in that plane (the thickness of a plane divided by the cosine of the incident angle of the shower). The transverse dE/dx is calculated using the following method. (1) A line connecting the start and end point of the shower is constructed. (2) In each plane the cell that the line intersects is considered to be the core of the shower and is assigned a transverse cell index of zero. (3) In a given plane, the next cells out from the core cell in the positive or negative transverse direction are both assigned a transverse cell index of 1, and so on for transverse cell indices up to 20. (4) For a given transverse cell index, the energies in cells along the entire shower are summed and divided by the total

path length to give the total dE/dx corresponding to that transverse cell index. (5) The average dE/dx for each transverse cell index is calculated.

By matching the reconstructed shower direction to truth, we select Monte Carlo e , γ , μ , π^0 , p , n , and π^\pm showers from neutrino MC events to extract the expected dE/dx distribution histograms for each plane and each transverse cell index. To consider energy dependence, we evenly divide the shower energy range 0-5.0 GeV into 10 bins, then obtain dE/dx histograms in these energy bins. Figures 1 and 2 show the expected longitudinal dE/dx distributions in different planes for electrons, photons and muons with energy greater than 0.5 GeV. In the second plane (plane index = 1), as shown in Figure 1(left), the electron has a sharp minimum ionization peak, while the photon has started the EM shower which has a broader distribution in dE/dx . This signature provides powerful separation of electrons from γ 's or π^0 's. Figure 2 (right) shows that in the 11th plane the electron is a shower whereas the muon still has minimum ionizing deposition, which makes the dE/dx distributions very different for these two particles.

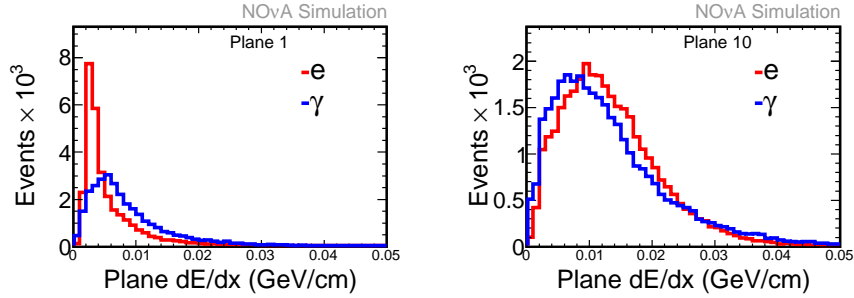


Figure 1: Longitudinal dE/dx for electrons (red) and photons (blue): (left) plane index = 1; (right) plane index = 10.

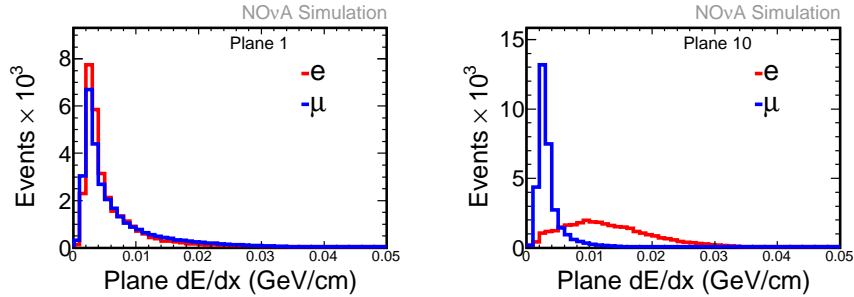


Figure 2: Longitudinal dE/dx for electrons (red) and muons (blue): (left) plane index = 1; (right) plane index = 10.

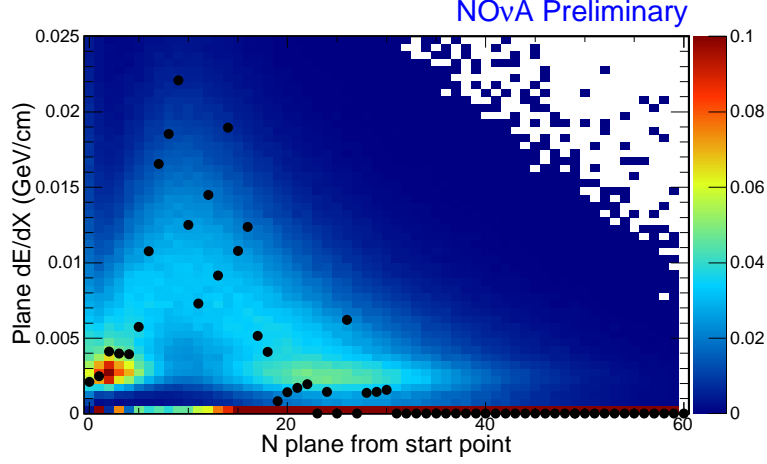


Figure 3: Color: p.d.f. for dE/dx in each plane (electron assumption); Points: measured dE/dx in each plane (example event)

When performing EID on the test sample, we calculate the dE/dx in each plane and transverse cell index. We then calculate the probability for each type of particle according to the expected dE/dx histogram for that plane or transverse cell index. We find the bin index in the expected dE/dx histogram corresponding to the test sample's dE/dx , and count the number of entries in that bin in the expected dE/dx histogram (N). The probability in plane (transverse cell) i is calculated as $P_i = \frac{N \times n_{bin}}{N_{tot}}$, where n_{bin} is the number of bins in the dE/dx histogram, and N_{tot} is the total number of entries in that histogram. The scale factor of n_{bin} is applied in order to avoid issues with the limit of machine precision when we sum $\ln(P_i)$ over planes or transverse cell indices to calculate the likelihood. Figure 3 shows the measured plane dE/dx of a signal-like event in FD superposed on the probability density as a function of dE/dx in each plane created by electrons in the FD signal neutrino MC.

The likelihood of each particle's hypothesis in the i -th plane (transverse cell index) is defined as: $LL_i = \ln(P_i)$. The overall longitudinal and transverse log likelihoods are defined as: $LL_L = \sum LL_i / N_p$ and $LL_T = \sum LL_i / N_t$, where N_p is the number of planes and N_t is the number of transverse cells in the shower. Differences between longitudinal and transverse log likelihoods for the electron and other particle hypotheses can be used to identify electrons. Differences between longitudinal and transverse log likelihoods for electron and other particle hypotheses can be used to identify electrons. As examples, Figures 4 and 5 show the difference of the electron likelihood and the likelihood of the μ/π^0 hypotheses in Monte Carlo neutrino events. These distributions illustrate the discrimination power against major backgrounds.

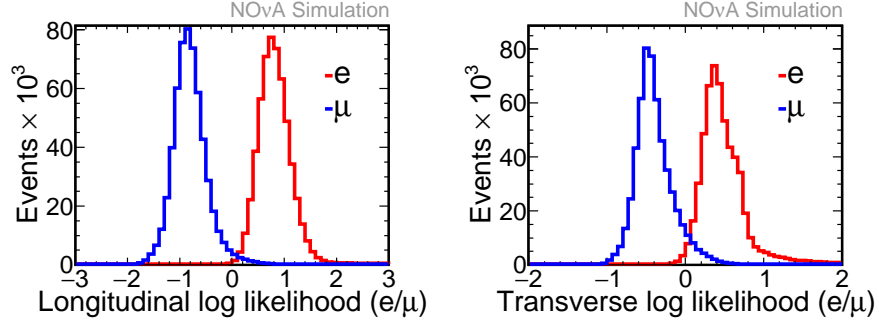


Figure 4: (left) Longitudinal e^- likelihood minus μ likelihood for e^- (red) and μ (blue); (right) Transverse e^- likelihood minus μ likelihood for e^- (red) and μ (blue).

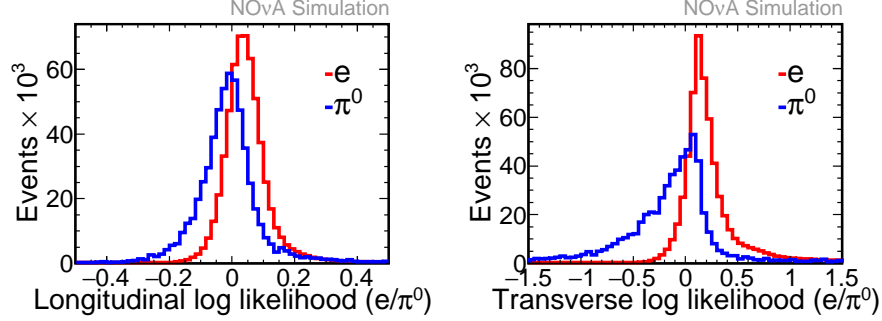


Figure 5: (left) Longitudinal e^- likelihood minus π^0 likelihood for e^- (red) and π^0 (blue); (right) Transverse e^- likelihood minus π^0 likelihood for e^- (red) and π^0 (blue).

2.5 Input variables and training of the artificial neural network

Seventeen variables are used to form the input of an artificial feed-forward neural network (ANN) to produce the likelihood-based ν_e identification variable - LID. The first category of inputs consists of 12 differences between longitudinal and transverse log likelihoods for electron and the six other-particle hypotheses (γ , μ , π^0 , p , n and charged π) applied to the most energetic shower [$LL_L(e) - LL_L(\gamma)$, $LL_T(e) - LL_T(\gamma)$, etc]. In addition to these log likelihoods, 5 additional inputs are used: (13) π^0 mass, the invariant mass of the most energetic shower with all the other showers in the slice is computed and the one closest to the π^0 mass is recorded to help reject π^0 s in neutral current interactions; (14) shower energy fraction, shower energy divided by the total event energy; (15) vertex energy, the calorimetric energy, excluding the leading shower, within ± 8 planes of the event vertex; (16) gap, the distance of the start point of the shower from the event vertex, which is small for electrons and large

for photons; $(17) \cos \theta$, angle of the leading shower with respect to the beam direction, for rejection of neutral-current interactions and cosmic rays.

The neural network has one input layer, three hidden layers and one output layer. The input layer has 17 nodes which correspond to the 17 input variables. The output layer has one node. The optimal architecture of the three hidden layers was determined to be 22:12:6.

2.6 Performance and choice of PID

The output of the artificial neural network (LID) can be found in Figure 6. The dashed line and arrow show the cut applied to the first ν_e analysis. The value, $LID > 0.95$, is obtained by maximizing the figure of merit defined as $FOM = S/\sqrt{B}$, where S is the number of signal and B is the number of the background. The intention of using $FOM = S/\sqrt{B}$ instead of $FOM = S/\sqrt{S+B}$ is to optimize the ability to reject the zero $\nu_\mu \rightarrow \nu_e$ hypothesis. The ν_e selection efficiency achieves an efficiency of 34% relative to the contained sample, while rejecting more than 99% of beam backgrounds.

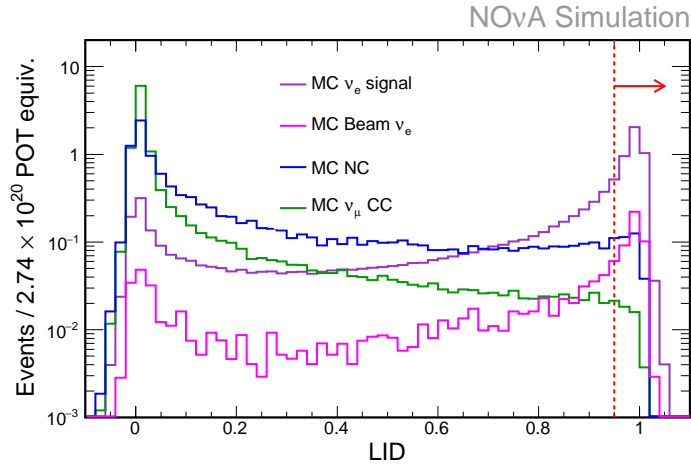


Figure 6: LID on FD neutrino MC (log. scale).

Similarly to LID, the cut for selecting ν_e events with LEM is determined to be $LEM > 0.8$. The ν_e selection efficiency is 36% relative to the contained sample, while rejecting more than 99% of beam backgrounds. There is 62% overlap of selected signal events between LID and LEM.

Both EIDs are very similar in signal efficiency, figure of merit (FOM), systematic uncertainties, and overall sensitivity to ν_e appearance and oscillation parameters. Prior to unblinding we decided to present results of both selections and to use the LID technique as the primary result.

3 Event selection and FD prediction

A series of data quality cuts are applied to ensure that the data subjected to the full analysis were taken under the normal beam and detector conditions. Events in a 12- μ s window around the 10- μ s neutrino spill time are selected to reject cosmic rays collected in the 550- μ s trigger window. Prior to the PID selection, reconstruction quality and containment cuts are applied to remove particles that are poorly reconstructed or that enter from the edges of the detector. An additional cut on the ratio of the transverse momentum to the total momentum of the event (P_T/P) is applied to require the orientation of showers to match the beam direction. The oscillation maximum is around 2 GeV. To further improve the figure of merit, the slice calorimetric energy E is required to be $1.5 \text{ GeV} < E < 2.7 \text{ GeV}$ ($1.3 \text{ GeV} < E < 2.7 \text{ GeV}$ for LEM). Finally, the LID value is required to be greater than 0.95 (0.8 for LEM) to select ν_e CC candidates. After all selection cuts, both EIDs have an overall rejection rate of 1 in 10^8 for cosmogenic backgrounds. Based on the cosmic ray data, we predict 0.06 cosmic background events for both LID and LEM [14]. The consistency between data and MC has been validated based on the near detector data and cosmic ray data [14, 15, 16, 17].

The Near Detector data provide a data-driven correction for the MC normalization in the Far Detector. We scale up each background component (ν_μ CC, NC and beam ν_e CC) in the Far Detector MC by the extrapolated data/MC ratio in the Near Detector to make the background prediction. Because there are no oscillated ν_e CC events in the Near Detector, we correct the Far Detector ν_e CC signal by the extrapolated data/MC difference of selected ν_μ CC events in the Near Detector. The Near-to-Far extrapolation ratios for signal and backgrounds are determined with Monte Carlo samples. This extrapolation cancels most of the systematic errors. Remaining systematic uncertainties after the extrapolation are evaluated by extrapolating ND data with nominal MC and systematically modified MC samples, with variations to the calibration, non-linearity in detector responses, the neutrino-interaction model, the ND background composition, overall normalization, beam uncertainties, and other smaller uncertainties. Summed in quadrature the total uncertainty is 17.6% for the signal and 10.1% for the background under the LID selection [16, 17].

Based on the Near Detector data and cosmic ray data, the expected background in FD is 0.94 ± 0.10 (syst.) events for the LID selector. Oscillation parameters that produce this prediction are $\sin^2 \theta_{23} = 0.5$, $\Delta m_{32}^2 = +2.37 \times 10^{-3} \text{eV}^2$, $\sin^2 2\theta_{12} = 0.846$, $\Delta m_{21}^2 = 7.53 \times 10^{-5} \text{eV}^2$, $\sin^2 2\theta_{13} = 0.086$, and $\delta_{CP} = 0$. The background prediction varies about %1 for different choices of oscillation parameters. The dominant backgrounds are 49% beam ν_e CC and 38% NC. ν_μ CC and cosmic rays are each about 6%. For the LEM selector, we predict to have 1.00 ± 0.12 (syst.) background events with 46% beam ν_e CC and 40% NC in them. ν_μ CC and cosmic rays are at the same level as LID.

The signal prediction depends on the choice of oscillation parameters. The highest prediction is made under the assumption of normal mass hierarchy, δ_{CP} close to $3\pi/2$ and $\theta_{23} = \pi/4$, where we expect 5.62 ± 0.99 $\nu_\mu \rightarrow \nu_e$ oscillation events selected by LID and 5.91 ± 0.89 $\nu_\mu \rightarrow \nu_e$ oscillation events selected by LEM. Before unblinding, we checked sidebands of EIDs and energy. Results of both are consistent with expectations.

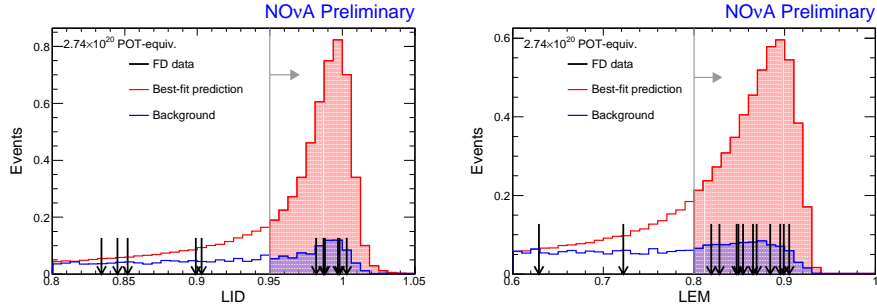


Figure 7: PID distributions for FD data and MC: (left) LID (right) LEM. Gray lines and arrows are PID cuts and selected signal region

4 Results

In the Far Detector data, 6 events are selected by $LID > 0.95$ and 11 events are selected by $LEM > 0.8$, as shown in Figure 7. In the data the (LID only)/(LEM only)/(LID and LEM) events are 0/5/6. The 5 events that are only selected by LEM have LID values in the range $0.7 - 0.95$, which is just below the signal region $LID > 0.95$. As shown in Figure 7 (left), in this LID region the probability of signal (red line) is still higher than background (blue line). In addition, 2 of the 5 LEM-only events also fail the energy cut $1.5 - 2.7$ GeV for LID. Given the expected correlations, (LID only)/(LEM only)/(LID and LEM) events in MC signal and backgrounds, the observed event counts yield a reasonable mutual p-value of 10%.

The ν_e appearance significance based on the 6 ν_e candidates selected by LID is 3.3σ (5.4σ for LEM). Allowed regions of δ_{CP} and $\sin^2 2\theta_{13}$ are obtained by the Feldman-Cousins (FC) approach to interpret the measured $-2\Delta\log(L)$ in the data [18]. When generating pseudo-experiments at each $(\delta_{CP}, \sin^2 2\theta_{13})$ point, central values and variations of other oscillation parameters are listed below: $\Delta m_{21}^2 = 7.53 \pm 0.18 \times 10^{-5} \text{ eV}^2$ and $\sin^2 2\theta_{12} = 0.846 \pm 0.021$ are taken from PDG [19], $\Delta m_{32}^2 = 2.37 \pm 0.15 \times 10^{-3} \text{ eV}^2$ [normal mass hierarchy (NH)], $-2.40 \pm 0.15 \times 10^{-3} \text{ eV}^2$ [inverted mass hierarchy (IH)] is taken from the first NOvA ν_μ disappearance measurement that is parallel to this analysis, $\sin^2 \theta_{23}$ is held at 0.5. The resulting allowed regions produced by LID

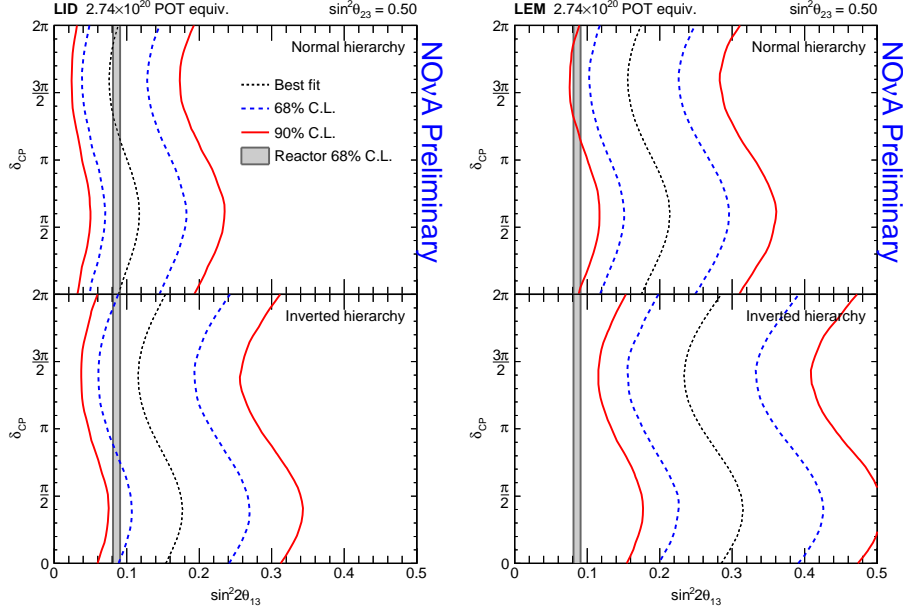


Figure 8: The allowed regions of δ_{CP} and $\sin^2 2\theta_{13}$ with the LID (left) and LEM (right) analysis. World average reactor $\sin^2 2\theta_{13}$ results are drawn as the gray bands.

and LEM are shown in Figure 8. World average reactor result $\sin^2 2\theta_{13} = 0.086 \pm 0.05$ are also shown as grey bands. For LID, the 68% C.L. interval for the NH hypothesis is compatible with the reactor result for all δ_{CP} values.

By varying $\sin^2 2\theta_{13}$ values within the global reactor constraint, we converted the 2-D FC contours into significance as a function of δ_{CP} , as shown in Figure 9. The LID result [Figure 9 (left)] shows that the range of $0 < \delta_{CP} < 0.65\pi$ in the IH is disfavored at the 90% C.L. in the case of $\sin^2 \theta_{23} = 0.5$. The number of events selected by LEM is greater than that predicted even in the NH, $\delta_{CP} = 3/2\pi$ case, but 12% of pseudo-experiments generated in this scenario have an equal or worse χ^2 than that observed in the data. The LEM result [Figure 9(left)] shows that IH is disfavored at $> 2.2\sigma$ while NH for $0 < \delta_{CP} < \pi$ is mildly disfavored ($> 1\sigma$). Both the LID and LEM results prefer normal ordering in most of the δ_{CP} range, and prefer δ_{CP} near $3\pi/2$.

References

- [1] D. S. Ayres et al. (NOvA), FERMILAB-DESIGN-2007-01 (2007); R. B. Patterson, for NOvA, Nucl. Phys. Proc. Suppl. 235-236, 151 (2013)

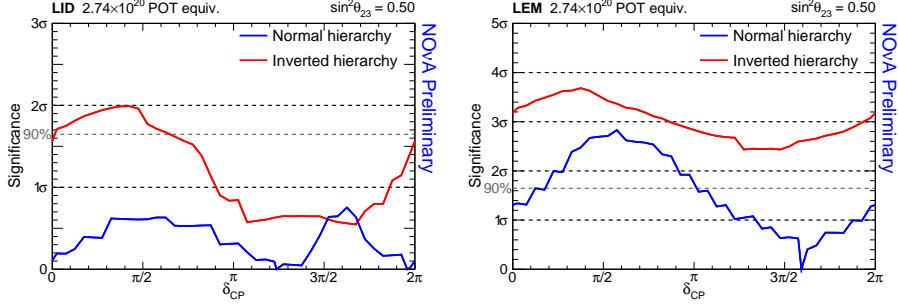


Figure 9: Significance as a function of δ_{CP} : (left) LID (right) LEM.

- [2] NuMI Technical Design Handbook, NuMI Technical Design Handbook, http://www-nuui.fnal.gov/numwork/tdh/tdh_index.html
- [3] S. Magill, J. Phys.: Conf. Ser. 404, 012035 (2012).
- [4] S. Mufson et al., Nucl. Instrum. Meth. A **799** (2015).
- [5] J. J. Grudzinski et al., FERMILAB-PUB-14-391-PPD, 2014.
- [6] P. Adamson, *et al.* [The NOvA Collaboration], “First measurement of electron-neutrino appearance in NOvA,” FERMILAB-PUB-15-262-ND, to be submitted to Phys. Rev. Lett.
- [7] C. Backhouse and R. B. Patterson, “Library Event Matching event classification algorithm for electron neutrino interactions in the NOvA detectors,” Nucl. Instrum. Meth. A **778**, 31 (2015) [arXiv:1501.00968 [physics.ins-det]].
- [8] M. D. Baird, “An Analysis of Muon Neutrino Disappearance from the NuMI Beam Using an Optimal Track Fitter,” Ph.D. Thesis, FERMILAB-THESIS-2015-24.
- [9] M. Ester et al., Proc. of 2nd International Conference on Knowledge Discovery and Knowledge Engineering and Knowledge Management, pp226-231 (1996).
- [10] L. Fernandes and M. Oliveira, Patt. Rec. 41, pp299-314, (2008).
- [11] M. Gyulassy and M. Harlander, Computer Physics Communications 66, 31-46 (1991); M. Ohlsson, and C. Peterson, Computer Physics Communications 71, 77-98 (1992); M. Ohlsson, Computer Physics Communications 77, 19-32 (1993); R. Fruhwirth and A. Strandlie, Computer Physics Communications 120, 197-214 (1999).

- [12] R. Krishnapuram and J. M. Keller, IEEE Trans. Fuzzy Syst. 1 (1993); M. S. Yang, and K. L. Wu, Pattern Recognition, 39 (2006).
- [13] E. D. Niner, “Observation of Electron Neutrino Appearance in the NuMI Beam with the NO ν A Experiment,” Ph.D. Thesis, FERMILAB-THESIS-2015-16.
- [14] T. Xin, Ph.D. Thesis, Iowa State University (2016).
- [15] H. Duyang, “Cosmic Ray Induced EM Showers in the NO ν A Detectors,” arXiv:1511.00351 [physics.ins-det].
- [16] K. Sachdev, “Muon Neutrino to Electron Neutrino Oscillation in NO ν A,” Ph.D. Thesis, FERMILAB-THESIS-2015-20.
- [17] J. Liu, Ph.D. Thesis, College of William and Mary 393 (2016).
- [18] G. J. Feldman and R. D. Cousins, “A Unified approach to the classical statistical analysis of small signals,” Phys. Rev. D **57**, 3873 (1998) [physics/9711021 [physics.data-an]].
- [19] K. A. Olive *et al.* [Particle Data Group Collaboration], Chin. Phys. C **38**, 090001 (2014).